

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

PETER A. JACKSON
M.A. (Soc Sci) E825.0650

Massey University Library

Thesis Copyright Form

Title of thesis: *The Dominance of the Physicalist-Reductionist Approach to the Study of Consciousness and its Evolution: The Case for a Non-Physicalist Paradigm*

(1) (a) I give permission for my thesis to be made available to readers in Massey University Library under conditions determined by the Librarian.

~~(b) I do not wish my thesis to be made available to readers without my written consent for ... months.~~

(2) (a) I agree that my thesis, or a copy, may be sent to another institution under conditions determined by the Librarian.

~~(b) I do not wish my thesis, or a copy, to be sent to another institution without my written consent for ... months.~~

(3) (a) I agree that my thesis may be copied for Library use.

~~(b) I do not wish my thesis to be copied for Library use for ... months.~~

Signed

Date

[Signature]

1 February 1993

PETER JACKSON

The copyright of this thesis belongs to the author. Readers must sign their name in the space below to show that they recognise this. They are asked to add their permanent address.

NAME AND ADDRESS

DATE

**The Dominance of the Physicalist-Reductionist Approach to the Study of
Consciousness and Its Evolution: The Case for a Non-physicalist Paradigm**

A thesis presented in partial
fulfilment of the requirements for the
degree of

Master of Arts in Psychology at Massey University

Peter Anthony Jackson

1992

To Kathy

for her love and enthusiasm
and for believing I could do it

ABSTRACT

In this thesis it is argued that the dominant physicalist-reductionist view in psychology has hindered the study of consciousness and its evolution. The thesis begins with an overview of the physicalist-reductionist position, from a philosophical viewpoint. The weakness of this position is exposed in that matter can be viewed not as a physical substance, but as the derivative of a non-physical realm. This is argued by using the theory of David Bohm, who postulated the existence of an implicate order (hidden from the senses) and an explicate order (the sensory realm). Bohm's reasoning is explored and justified, where his theory is shown to be the way to reconciling the difficulties faced by quantum mechanics and relativity theory. Using Bohm's implicate-explicate notion as a basis, an implicate-explicate model of the evolution of consciousness is developed. This model is able to deal with biological evolutionary factors, and not simply with the evolution of consciousness. This is because the model assumes that biological forms evolved as a result of the interaction between the implicate and explicate orders. In this model, evolution is seen as the unfolding of what lies implicate, which then becomes explicate. The earliest stages of this unfolding were automatic, and led to increasingly complex physical, chemical, then biological structures. A stage was reached where biological structures gave rise to sufficiently complex neural structures which, in turn, permitted consciousness to appear. In this model, consciousness is a very high order explicate of a special region within the implicate order, which is called Mind. Thus, the evolution of consciousness is the result of the unfolding of Mind. The model shows that consciousness is an active factor in the further evolution of biological forms. The notion of consciousness is explored and a variety of theories of consciousness are reviewed and critiqued, where these are examined in the light of the implicate-explicate model. This model is then used to explore the way consciousness evolves through the infrahuman life forms to the human form. Palaeoanthropological evidence is used to justify the claim that

consciousness has evolved, with a special focus on primate evolution, and on the critical phase of transition from proto-human to truly human consciousness. In this, the acquisition of speech is seen as crucial, where the implicate-explicate model offers an explanation for this acquisition. The notion of psychological paradigms is explored, and a set of paradigms delineated, where these are located along a spectrum of the relevance of consciousness to any given paradigm. The relevance spectrum is related to the implicate-explicate model as a metaparadigm. This is used to reveal the strengths and limitations of the various paradigms. The implicate-explicate model shows that present-day humans have reached an impasse in the evolution of their consciousness. A means to overcoming this is suggested, and the next stage in the evolution of consciousness that might arise is speculated upon.

PREFACE

The topic of consciousness and its evolution has long fascinated me. This fascination was a major reason for studying psychology at the undergraduate and graduate levels, where I hoped to find some answers to the mystery of consciousness and its evolution. While I found answers to many other things about human nature and behaviour, I found psychology (with a few notable exceptions) had little to say on the topic of the evolution of consciousness. As far back as the 100 level papers, the absence of the *psyche* in psychology concerned me. This concern was aggravated by the relative silence in psychology regarding the evolution of consciousness and the dominance within this discipline of physicalist-reductionist thinking. I had, perhaps naively, thought that psychology was the one viable agent for change. I had seen it as the one branch of science that could shed light on human nature, its purpose and its possible future. I have since had some of the naivety knocked out of me, but my faith in psychology as an agent for change lingers on.

A breakthrough in my study of consciousness came from my readings in a field far removed from psychology. This field was that of physics, quantum mechanics in particular, in which the consciousness of observers plays a crucial role. Having given the paradoxes of quantum mechanics some considerable thought, and having been deeply impressed by the work of David Bohm, a theoretical physicist, I began to see why the topic of consciousness was being avoided by psychology, but also saw the possibility of a way forward to a better understanding of it. Moreover, in the work of Bohm, I saw the basis of the possibility of developing a model of consciousness and its evolution. Doing a special 400-level topic on the evolution of consciousness with Dr Dave Clarke

(Department of Psychology, Massey University) firmly cemented my interest in this subject where, during that course, the idea to write this thesis arose.

The remnant of my faith in psychology as an engine for change has led me, perhaps egotistically, to believe that an area of contribution that I could make would be to edge the psyche back into psychology, even if through the back door. This prompted me to write this thesis in the hope that it might provoke thought and suggest possible ways forward to a wider interest in and deeper understanding of consciousness and its evolution. Were this hope even vaguely realised, then one small contribution to a needed paradigm shift might have been made.

ACKNOWLEDGEMENTS

I wish to thank my two thesis supervisors, Dr Dave Clarke (Psychology) and Dr Tom Bestor (Philosophy), for all the effort they have put into to keeping me on the rails. Dave's enthusiasm for this project inspired me to write it, and kept me going when my motivation hit rock-bottom. I have also leaned on Dave's deep grasp of human nature and his very wide knowledge of psychology, where his guidance on the psychological arena has been invaluable. Tom's clarity of thought and logical analysis were vital in a thesis that is as much philosophical as it is psychological. Tom does not pull his punches. I have valued that even when the punches were painful, and have learned something of the self-control that is needed to write a thesis. With wisdom and compassion, Tom clipped my wings often, but he also let me flutter about at times. At least now, I have some idea as to what it might be like to fly. Thank you Dave and Tom.

I wish also to acknowledge my indebtedness to Professor David Bohm, without whose years of brilliant work in the field of theoretical physics, my thesis would have had no empirical basis. I believe that Bohm's ideas are as far ahead of their time today, as were Einstein's when the Newtonian paradigm still held sway. I have widely used (and perhaps abused) his ideas, and pushed them into areas he might not approve of.

I thank Ian Timperly for the highly professional art-work he executed in the illustrations used in this thesis.

I also want to thank Miles, my ten-year old son, for all his patience and understanding during those times when his dad locked himself away in his study.

Most of all, I want to thank Kathy, the light of my life, whose faith in me kept me going during some rather dark moments. She listened with interest to my ramblings, lifted me out of depression, helped me break my log-jams and has read every line of this fairly long thesis.

CONTENTS

Preface	(v)
Acknowledgements	(vii)
CHAPTER 1: INTRODUCTION	i
A brief history of philosophical physicalism	1
The weakness of the physicalist case	8
The notion of evolution	10
Consciousness	14
The implicate-explicate set of categories	21
The evolution of consciousness	23
CHAPTER 2: THE IMPLICATE AND EXPLICATE ORDERS	25
The nature of physical reality	25
Quantum mechanics	26
Four key questions or theses	33
David Bohm on physical reality	36
A new theory of reality	42
The ink-in glycerine metaphor	43
The holographic metaphor	47
The Bohm-Pribram misconception	54
The relationship between implicate and explicate	55
The mind, consciousness, brain relationship	57

CHAPTER 3: THE IMPLICATE, EXPLICATE, CONSCIOUSNESS AND EVOLUTION	60
The concept of evolution	61
A clockwork analogy	64
Feedback from explicate to implicate	66
A new clockwork analogy and the notion of shape	69
Death as a feedback mechanism	73
A simple computer analogy	77
A summary to this point	82
The entry of consciousness	83
Computer-based analogies	88
The influence of Mind at the explicate level	96
A summary of the modes of unfolding	100
The notion of implicate regions	104
A dual evolutionary process	107
The unfolding of the implicate order	109
Consciousness as an active agent	114
 CHAPTER 4: THE SPECTRUM OF CONSCIOUSNESS AND ITS EVOLUTION	 118
Theories of the evolution of consciousness	119
Intellectualist theories	123
a sociobiological bias	123
art and the evolution of consciousness	125
Holistic theories	128
the self and its brain	128

a dualistic model	130
the role of speech	133
the bicameral mind	138
a multidimensional view	140
A summary: Current theories of the evolution of consciousness	145
Current theories in the light of the I-E model: A summary	148
 CHAPTER 5: THE EVOLUTION OF PRIMATE CONSCIOUSNESS	154
Primate evolution	154
geological eras	154
pre-hominid evolution	155
the primate family tree	158
hominid evolution	159
The kinds of evidence supporting the evolution of primate consciousness	163
The implicate-explicate model and primate evolution	173
The evolution of consciousness: where to next?	190
 CHAPTER 6: THE IMPLICATE-EXPLICATE MODEL: SOME IMPLICATIONS FOR PSYCHOLOGY AND FOR THE FURTHER EVOLUTION OF CONSCIOUSNESS	193
The notion of a paradigm	193
The major psychological paradigms	194
The spectrum of paradigms	200
The implicate-explicate model revisited	204
Psychological paradigms and the implicate-explicate orders	207

The evolution of consciousness: an epilogue	211
restarting the evolution of consciousness	212
the next step in the evolution of consciousness	216

GLOSSARY OF ABBREVIATIONS

A. afarensis: Australopithecus afarensis

A. africanus: Australopithecus africanus

AI: Artificial Intelligence

BP: Before the present

CAI: Computer Assisted Instruction

CNS: Central Nervous System

E: Explicate Order

E-mail: Electronic mail

He: Homo erectus

Hh: Homo habilus

Hs: Homo sapiens

Hsn: Homo sapiens neanderthalensis

Hss: Homo sapiens sapiens

I: Implicate Order

I-E: Implicate-Explicate

LA: Law of Approximation to an Ideal

LC: Law of Increasing Complexity

LE: Local Environment

LU: Law of Unfolding

S: Shape (at birth)

S': Shape (at death)

LIST OF FIGURES

Figure 2.1: Producing a hologram	50
Figure 2.2: Retrieving the stored holographic image	51
Figure 2.3: The hologram as a metaphor of the relationship between consciousness and the implicate order	56
Figure 3.1: Relationship between the various analogies used and the implicate(I)-explicate(E) modes	101
Figure 3.2: The relationship between the implicate and explicate orders in the progressive unfolding of the implicate	110
Figure 6.1: The psychological paradigms, consciousness and the spectrum of explication	203

CHAPTER 1: INTRODUCTION

This thesis addresses itself to what I regard as the extreme imbalance that exists in the current scientific view of consciousness and its evolution. The imbalance is created by a strong bias toward a physicalist-reductionist view of consciousness. This places a severe constraint on the view of the nature of consciousness, and rules out any other view or explanation. In this context, the compound term *physicalist-reductionist* describes a philosophical position which asserts that there is only matter in some state or another (it may be fields, subatomic particles, or tangible substance) and that, in the final analysis, all animal behaviour (including that of humans and their societies) is reducible to the laws that govern matter (usually thought to be that of classical physics). This physicalist-reductionist account has served well the physical sciences (such as, physics, geology and astronomy) in the sense that they have achieved a high degree of prediction and control. However, I argue that this view is misplaced in the human sciences (eg, psychology and sociology).

A BRIEF HISTORY OF PHILOSOPHICAL PHYSICALISM

In fairness to the viewpoint I call *physicalist-reductionist*, its development and basic tenets should be made available to my reader, so that an informed judgement can be made in regard to the validity of the arguments of both sides. To this end I give a brief history of the development of physicalism.

In the development of the physicalist-reductionist viewpoint, we can identify three separate stages: Identity Theory, Functionalism and Eliminativism.

Identity theory

The attractiveness of Physicalism is that it avoids the difficulties inherent in having a dualistic scheme where there are mental things and physical things. That is, it makes life very simple if every thing in the world, including humans and their complex behaviours, can be reduced to one ultimate class of things (matter of some form). Put more formally, this argument goes that there is nothing we can say about minds that is not already exhausted in

claims about the central nervous system. In this approach we are saying that there is only the one thing, physical matter, which we refer to accurately with physical terms and which we refer to misleadingly with mentalistic terms. That is, the essential Physicalist claim is that there is no other thing but physical matter, which mental terms refer to.

The vindication of this claim was, and still is, the goal of Physicalism, and it strongly motivated the early Physicalists. In pursuing this goal, the early Physicalists discovered (perhaps unearthed) what has become known as *Leibniz Law*. This Law was not formally stated by Leibniz, but was implicit in his mathematico-philosophical writings. The Law can be stated as: *if one thing is identical to another then anything that is true of the one must also be true of the other*.

The great value of Leibniz Law is that it makes it very clear that if we wish to equate two separate things with an equals sign, then the property lists of each of the two things must be identical. That is, even though we are using different labels for the two things (say, x and y), they are identical if the listings of their individual properties are identical. If this is true, then we can validly say $x = y$.

It was the application of this line of reasoning that led Descartes to his dualistic view of the mind-body problem. He said that the properties of minds are quite obviously different from those of the central nervous system (eg, the sensation of pain is quite different to the firing of certain clusters of neurons). Descartes said that mental states are conscious, private, non-spatial and indivisible, whereas physical states are non-conscious, public, spatial and divisible. When viewed in this way, there can be no Leibnizian identity because the property lists are clearly different.

The early Physicalists saw the truth in Descartes basic argument. They realised that if one applied Leibniz Law down at the level of specific sensations and given clusters of neurons then, according to that law, their case was hopeless. They saw that the power of Leibniz Law

lay in providing the means of checking for a true identity, and that if they were to obtain this condition in regard to the mind-body issue, they needed to look above the level of specific sensations and specific neurons. They saw that their earlier claim that "The mind is nothing over and above the central nervous system" was an unsafe route to follow. In fact, it became obvious that the only way to meet the Leibnizian Identity test was at the level of an entire person. Thus, they saw that the property lists were identical only when one said that a person experiencing some state (x) had a brain that was in some state (y). In this, one can legitimately say that $x = y$.

However, it was this perfectly logical application of Leibniz Law by these early Physicalists that led to a crisis in Identity Theory. While Leibniz Law seemed to have come to the rescue of Physicalism, it also enabled the opposition to formulate objections to the classical notion of identity. In fact, three major objections to classical identity theory were highlighted. The first of these had been effectively raised by Descartes, who said that the property list of mental things was not identical to that of physical things. To give them credit, as stated above, the Physicalists saw the truth in this, and worked to overcome it. The other two major objections were not so readily overcome. The first is that humans have a strong sense of the privacy of their minds, and that their minds are their own. This is known as the *first person* objection, where Classical Identity Theory leaves this something out. The second objection is that Identity Theory is species-chauvinistic in that it insists that only creatures with brains can have mental states, because it insists that mental states and brains are identical.

Taken far enough, this final objection to identity theory not only insists that other forms of life (eg non carbon-based) could have mental states, but that we cannot even restrict mental states to structures per se. The way out is to identify a given mental state with a functional state, which state could be realised in any number of structures or non-structures. This final objection to classical identity theory led on to Functionalism.

Functionalism

There are three key ideas underlying the functionalist view. Firstly, there is a distinct difference between a functional description and a structural description. In computer science, this difference is captured by the difference between, say, a hardware circuit diagram of the central processor and a listing of the program code. The latter says nothing about the structures it might use to get things done, and the former says nothing about the processes that might occur within the described structures. Secondly, the functional descriptions can be hierarchically arranged, with lower-order functions nested within higher-order ones. For example, the organisational chart of a company is hierarchical in that it shows the Chief Executive Officer at the top, various senior executives at the next level down who head certain functions such as corporate services and human resources, then levels of middle management, and finally those who report directly to these managers. The third idea is that, if the lowest level of such a hierarchy consists of elements so simple that they are purely mechanistic, then physicalism is vindicated.

Combining these ideas enables us to correlate mental stuff with the functioning of the whole. In this scheme, identity statements are still relevant, but become more subtle. In this case, where we are equating a functional device (say a text manipulation device) with a structural device (say, a personal computer), the function (text manipulation) can be carried out by a number of different structures (knife and woodblocks, ball-point, paper and scissors, and a personal computer using wordprocessing software). That is, we are saying in our *identity* statement that in this case that specific function is carried out by this specific structure, thus: *a personal computer is a symbol cruncher*.

When we get to human examples, it gets even more complex because, in the Cartesian tradition, the structures that have the function of, say, concept-maker or idea-abstracter are mental structures existing in a mental space. But, the functionalist disagrees, and argues that, in the end, there is an identity between mental states (functions) and physical structures (eg, a human brain). That is, these specific functions of abstracting concepts, say, is carried out by the specific structure of a neocortex, thus: *thinking is the firing of neurons in the neocortex*.

While classical functionalism was a great advance on classical identity theory, it still had problems. In particular, any mental state or process has a subjective feel to it (see the first person objection), and it's hard to see where a subjective element comes into the structuralist story. For example, pain is not a good piece of structure, nor is it any better as a function. Such things as pain just happen and are phenomenal. That is, events do not figure in the function-structure model.

Another problem is that of beliefs and propositional attitudes (eg, worries, fears, hopes, regrets etc..). Here are three such propositional attitudes:

(a) While I am typing this, I worry that the right keys will be missed being hit by my fingers.

(b) While I am typing this, I worry that my fingers will miss hitting the right keys.

(c) While I am typing this, I worry that the keys will miss being hit by the right fingers.

The problem is that we know that (a) and (b) are the same, and that (b) and (c) are different propositional attitudes. That is, (a) and (b) tell the reader that my worry concerns getting my fingers to the keys I really want to get them to, whereas in (c) my worry is about the keys and their sense of loss. We know this ahead of any application of a rule of structural and or functional identity, and especially ahead of any neurophysiological research. We know they are different propositional attitudes (ie, different functional states) purely only on the basis of the *grammar of English*. But this is not how identity or non-identity of physical states or functions is supposed to be determined on any basically physicalist-reductionist stance. We are supposed to look at *physical things*, not at grammar, for such decisions.

This is a problem for physicalism in general. For identity theory, the problem is that we can determine too many identities of structural states non-empirically (ie, non-physically). For functionalism, we can determine too many identities of functional states non-empirically (ie, non-physically). And, as we shall see, for the next view, eliminativism, the problem is how could questions of syntax and semantics even get a grip, if all there are are purely physical neurological events and processes.

Eliminativism

Beyond functionalist physicalism, there arose an eliminativist version, usually called Eliminativist Materialism. This approach is based on the notion that science moves forward by ditching the old as it takes on the new (a dubious and simplistic view, as Kuhn, 1962, has shown). Thus, if there are only brains, and thereby no minds, then our ways of psychologically speaking (called *Folk Psychology*) are meaningless and should be dropped.

Basically, what the eliminativist is saying is that any attempt to retain the theoretical terms of the Cartesian theory of mind, or attempts to retain our everyday psychological ways of describing human behaviour, have no value. Rather than retain the earlier notions and then try to squeeze them into a workable identity equation (in which the Folk Psychological terms go on one side of the equation, and neurological terms go on the other side), the eliminativist argues that we do not need an identity equation in the first place, because it causes so much confusion.

This is more than simply changing the kinds of explanations that are offered in a new theory. It means changing the very phenomena and the ways of identifying these phenomena which the new theory explains. The eliminativist argues that with any new theory, we do not try to hang on to the previous ontological entities and processes. Instead, we give a new explanation of their properties and relationships. That is, we completely redescribe the very phenomena which we take the new theory to be explaining. The eliminativist believes that once neuroscience reaches full maturity, we will no longer think of the phenomena which are

genuinely there and which have a real explanation in any of the terms which we now use, as inherited from our outmoded theories.

To summarise, from the above discussion, we can see that Eliminativism is a way around trying to keep mentalistic expressions and yet ditch minds. You get rid of the lot because any attempt to hang on to the Cartesian mode (or the Folk Psychological mode) of expression has no meaning. In this approach, the notion of $x = y$ is dropped as irrelevant because we are dumping the old Cartesian referent and not identifying two existently real things as with an equals sign. That is, x has no real existence or properties that science can determine, so it is completely replaced by expressions of the new theory. Thus, there are no X s, only Y s. This gets around the identity issue in the Leibniz and functional senses. However, to take this view to its logical conclusion means being able to drop the old vocabulary and replace it with the new scientifically approved vocabulary. This is done, because the old modes are regarded as containing no ontological truths, and that we retain them only through linguistic stubbornness. However, the chunk of the language we would have to drop is so large that the exercise might become absurd.

Eliminativism has struggled with the notion of the privacy of thought and how practically one could clean up the language to eliminate Folk Psychology. From these respective struggles, a spectrum emerges. At one end of this spectrum elimination is valid and at the other functionalism is valid. In the battle between the two, much depends on the findings of neuroscience and cognitive psychology.

Contemporary developments

The contemporary physicalist scene has changed considerably from that of the 1950s to 1970s. Physicalists have learned from the lessons of classical identity theory and functionalism. Essentially, science doesn't work by either *reducing* individual x s to y s or *replacing* individual x s by y s. It works by reducing entire theories (where x s dominate) to new theories (where y s dominate). Similarly with replacement.

Moreover, the progression of reduction or replacement is a matter of degree and not an either-or choice. The notion that the old theory is just a subset of the new theory is incorrect, because it is never the old theory which is reduced-replaced but a corrected version of it. The extent of this correction lies on a spectrum ranging from minor revision (at the classical identification-reduction end) to extreme changes which leave the old theory far behind (at the eliminativist-replacement end). For example, the transition from Aristotle's physics to that of Newton's entailed replacement, whereas the change from Newton to Einstein entailed reduction.

The most fundamental question for Physicalism is the extent to which there is a fit between a *Mature Neuroscience* and Folk Psychology. The real issue is whether neuroscience organises the phenomena for which it offers explanations in a way that is radically different from the way Folk Psychology organises the phenomena which it explains. This could go either way. In one direction, it might be shown that there are strong parallels between the categories of Folk Psychology and the neurological activities uncovered by neuroscience (eg, the notion of the *just noticeable difference* between two very similar weights may correspond well with the firing of certain clusters of neurons in the cortex). In this case, the fit would be good and a fair degree of reduction could occur. Conversely, there may be no significant correspondence between the two modes of explanation, so reduction cannot occur, and the Physicalist will want to ditch the Folk Psychology (for example, the clinical descriptors used for the various types of depression may have no correspondence with neuronal activity). From the Physicalist view, Folk Psychology is not worth keeping for its explanatory power because it is too dualistic in its ontological commitments, but bits of it may be found worthwhile for its conceptual organisation.

THE WEAKNESS IN THE PHYSICALIST CASE

This entire historical development of physicalism has validity only if the root assumption is valid. This root assumption is that there is basically only one kind of substance in the

universe -- physical matter (whether particulate or wave-like). This root assumption lies back of the more overt assumptions made by physicalism: that mind (and all other mental states or processes) can be seen only in terms of observable behaviours; that only science can judge what is real and what is not; that the essential causes of behaviour are neurological and physiological states.

It is arguments at the level of these three overt assumptions which have led to all the fine interplay of logic described above. But, if the root assumption is successfully challenged these arguments come to nothing. Thus, rather than get caught up in this logic chopping, I shall take a different route to undermining the physicalist bias in this thesis, by going straight to what I see as the essential weakness in the Physicalist's argument, which can be argued as follows.

The physicalist insists that while in mind and matter we have seemingly two different things, in reality they are two different ways of referring to the same thing. I point out, however, that there is nothing in this argument that shows that the physicalist's "*same thing*" must be *matter*. It could equally well be mind, or something totally other than either mind or matter. However, the physicalist insists that this same thing be matter. This insistence does not follow from the preceding logic. The strength (or possible weakness) of my thesis rests on my argument that what the physicalist calls matter, and what he/she regards as the mechanistic laws that govern this so-called substance, are ultimately intangible and arise out of something that could just as well be labelled Mind.

At this point, it is useful to summarise the overall strategy to be taken in in this thesis. I will first argue that there is a real, justifiable and valid distinction between two qualitatively different orders of reality. These are the explicate order (sensory reality) and the implicate order (hidden and non-sensory), based on the work done by the theoretical physicist David Bohm (Bohm, 1980). I will then demonstrate that philosophical physicalism does not recognise the explicate-implicate distinction. In fact, more than this, physicalism is not aware

of the distinction. Logically, this prevents physicalists from discussing anything other than physical type things, and so dramatically restricts the explanatory power of their paradigm.

It will then be argued that what is called mind is an aspect of the implicate order, and that what is called consciousness is a very high order explicate of mind (as wholly implicate). It will be further argued that the brain (human or otherwise) is wholly explicate, and of a lower explicate order than that of consciousness. I will then argue that, because all of the physicalist arguments in relation to the mind-brain issue are based solely within what I categorise as the explicate order, there is no intellectual room within physicalism to distinguish between brains and an implicate order. That is, the physicalist is forced to explain everything to do with the brain and its correlates only in terms of physical matter. This leads to the logical conclusion that the physicalist is working within the wrong category of things to be explaining the nature of mind, and hence the nature of consciousness. This category error prevents an adequate understanding and explanation of the relationship between brains, consciousness and mind. That is, while the brain is a necessary condition for an explanation of consciousness and its evolution, it is not a sufficient condition, as the physicalist maintains. This is because physicalism is ignoring the implicate order out of which the sensory world (explicate) emerges.

THE NOTION OF EVOLUTION

Because the two key terms in the title of my thesis are *evolution* and *consciousness*, it is necessary to spend some time considering these two terms and how they might come together in the phrase *evolution of consciousness*. This, of course, makes the assumption that consciousness is an entity in itself and that it does evolve. I do make this assumption and I regard the evolution of consciousness as a genuine phenomenon in nature.

When most people hear the word **evolution**, they think primarily of biological evolution and recall Charles Darwin (Darwin, 1859) as being the key figure in this field. However, the term evolution has a far wider meaning and application. I hold this view because the term

"evolution" is derived from the Latin *evolutio*, where its root meaning is that of the unrolling or unfolding of that which was wrapped or folded up. Biological evolution in the Darwinian sense has very little to do with the unfolding of something which was previously wrapped up. It has more to do with the increasing complexity of organic forms as a result of random mutation and natural selection. In saying this, I am not saying that evolutionary theory has it wrong. I am saying that biology made a poor choice of term in selecting the word "evolution".

I propose using the term "evolution" in two additional contexts, and for them I will thus give it a meaning quite different to that found in standard texts on biological evolution.

Firstly, I will use the term "evolution" in regard to *consciousness*, and for this special case give it a meaning of something unfolding, because, as I will show, the ***unfolding*** of that which was enfolded characterises the explicate order. I will later show that consciousness is a special explicate of the implicate order, and comes into being as a result of the unfolding of Mind (a very high region within the implicate order). This specific explication (of consciousness) can only occur once biological forms have reached a certain degree of unfolding, as I will show in a later chapter.

Secondly, I will use the term "evolution" in regard to *inorganic matter*. Assuming the Big Bang theory has validity, then primordial matter was far simpler than it is now, and went through some process to reach its current state. However, science restricts the application of the term "evolution" to biological forms, and does not apply it to the process whereby inorganic matter came to be as complex as it is. In my thesis, however, I will speculate that the process of unfolding the implicate order into explicate forms (the locus of David Bohm's use) applies equally well to subatomic, atomic, chemical and molecular structures. Here, I accept, I would be using the term evolution quite unusually. To avoid confusion, I will stay with the notion of an unfolding taking place within the implicate order. In this wider sense, I am referring to the development of matter from its most primordial state (the primordial singular atom, perhaps) to the present structure of the universe, which includes the atomic elements,

crystalline structures, planetary structures and the evolutionary sequence of stellar bodies. Indeed, in my view, it does not seem possible to consider biological evolution without conceding to some prior developmental process of atomic and molecular structures; and -- the point of my and Bohm's unifying category -- this developmental process is a process of unfolding the explicate from the implicate.

The distinction between non-living and living forms, and between pre-conscious and conscious forms, has never seemed very appropriate to me, in that it assumes that, at some critical level of complexity, what was previously dead matter suddenly becomes alive. If the matter that comprises cellular and other organic structures was not in some way already alive, I find it hard to see how it ever could become alive, simply by being in a more complex arrangement. Either subatomic matter is already alive in some sense, and therefore subject to some evolutionary process, or the notion of life at any level of complexity has little meaning. I speculate that what we term *Life* is inherent in all of nature, and is the manifestation of a law that operates in what Bohm (1980) calls the *holomovement* (to be dealt with in detail in Chapter 2).

In the human context, I shall define evolution as unfolding in terms of a change that is different from a purely quantitative increase (whether of individual cellular mass or numbers of humans) or individual growth-development across any given human life-span. Moreover, evolution is not necessarily a continuous process nor necessarily following some straight-line temporal law. At the elementary biological level evolution is a non-random or systematic change of gene frequency. For this reason, its effects manifest themselves in populations and not in individuals. Thus, in the biological sense, evolution seems to relate to population change, wherein a population is a collection of individual organisms. It is populations that evolve in the Darwinian sense and not individuals. However, this strictly Darwinian and reductionist biological view rejects other possibilities such as the notions of Lamarck (cited in Richards, 1987). In Chapter 3, I will explore some of these possibilities, including the

Lamarckian notion, and develop a model for the interaction between the implicate and explicate orders.

Earlier philosophical (eg, Henri Bergson), metaphysical and theological explanations of evolution have been in conflict with the Darwinian view, because they posited either an inner driving force (an **elan vital**) or some outer driving or guiding force (usually a divine teleological force). However, even thinkers with a strong metaphysical and spiritual bias, such as Pierre Teilhard de Chardin (Teilhard de Chardin, 1959) assume a Darwinian-like principle as external to the process. That is, even these thinkers do not seem to display enough faith in their inward-spiritual view, and tend to envisage evolution as a trial and error process which involves the generation of as large a variety as possible of attempts to proceed in multiple directions, where only the few of which are expected to breakthrough to higher states of evolution. This does not point to an inwardly driven process. Rather, despite giving credence to some mix of elan vital and teleological process, thinkers such as Teilhard are giving over a fair degree of control to chance-driven selection processes. It is only when we move away from Western influenced thinking to, say, Hindu metaphysical treatises, that we get closer to a truly inwardly driven evolutionary process (eg, Aurobindo, 1974).

It is worth pointing out that the Darwinian concept of biological evolution has been recently challenged, and that some theorists regard the whole theory as being in crisis, in that the typological view which preceded Charles Darwin, and which he argued against, is not to be so readily dismissed (Denton, 1985). The issue of typology will be taken up later, in relation to Bohm's (1980) theory of implicate and explicate orders, and how I see his notions providing explanatory power to the evolutionary concept. But, beyond any issue of criticising current evolutionary theory, I must address the issue of biological evolution because, in the view I wish to put forward in this thesis, the evolution of consciousness is inextricably bound up with that of biological evolution. Indeed, it is more appropriate to say that biological evolution is, in part, an outcome of the evolution of consciousness.

CONSCIOUSNESS

Before I deal directly with the term "consciousness" I wish to address the confusion that is liable to arise in the use of related terms such as "mind", "mental", "cognition", "awareness", "perception" and so on. For example, in some texts the terms "mind" and "consciousness" are used as synonyms. In my view, this leads to confusion, because these terms relate to two different things.

The word *mind* comes from the Anglo Saxon *gemynd*, which means *to think* or *to recall*, and is comparable with the Old High German word *gimunt*, which means memory. With thinkers such as John Locke, whose main philosophical work was published in 1690 (Locke, 1961) and Rene Descartes, who originally published his major work in 1641 (Anscombe & Geach, 1954), the word mind took on a deeper and more philosophical significance, with a meaning closer to the term soul or spirit. In fact, Descartes used *esprit* where he was referring specifically to what he regarded as the seat of reason (as residing in mental substance, where the pineal gland was the location of the causal interface between mind and body). Although Locke disputed the *a priori* nature of knowledge, and posited experience as the only given, he gave credence to a mind (an inner something) which dealt with what was acquired by the senses. We get another clue to the meaning of the word mind when we see that its earliest origin is in the Sanskrit term *manas*, which is the egoic principle or higher self in humans. From this, I assume that mind refers to something like an inmost self, as opposed to the every day waking self.

Not wishing to push the etymological issues too far, it seems clear that the term "consciousness" was derived from the Latin *con-scientia* (literally, with knowledge), where this term relates to inside knowledge (as opposed to knowledge of things external as in *scientia*). In this sense, consciousness seems to refer to a subjective entity which lies back of *scientia*. In my view, consciousness is synonymous with the word *self*. I here wish to distinguish between a *self* and a *higher self*, and so regard mind (higher self) as distinguishable from consciousness (self). I shall make this distinction clearer as I proceed.

Modern psychology has been uncomfortable with any notion of a **self**, because psychology has clung to objectivity, despite the highly subjective nature of its field of enquiry. The truth is, that consciousness has no particular objects of its own, because (at least in space-time) the objects of conscious experience are the same as those of non-conscious experience. To be conscious is to *live* the uniqueness of one's experience while transforming it into the universality of one's knowledge. Viewed in this way, we do not conflate conscientia and scientia, and clearly see that the latter is a function of the former.

I wish now to make certain key distinctions between terms such as mind, consciousness, awareness and cognition, because I shall use these terms in a special way that is different to their conventional usage in, psychology. Additionally, I need also to briefly discuss the issue of the relationship between consciousness (and its subordinate terms) and the brain.

1. Mind and consciousness: Taking into account the two quite different meanings I have given to the terms mind and consciousness (but without relying on etymological arguments), I shall argue that consciousness (as the everyday self) is a derivative of mind (the higher self), and is thus junior to it in a hierarchical sense. I will argue this not on the basis of etymology, but by using David Bohm's implicate-explicate theory as the basis for a model in which Mind is implicate, whereas consciousness is an explicate of Mind. Note that, from here on, I use Mind (with a capital M) rather than mind, so as to distinguish Mind (as a region within the implicate order) from its non-technical usage as in common discussion (eg, I have just made up my mind), from its usage as in philosophical arguments (eg those of the physicalist on the mind-body problem), and from its usage in most psychological paradigms (eg, as in cognitive science where it discusses the mind's representations of reality).

2. Consciousness and awareness: The term *awareness* is sometimes used as a synonym for consciousness. However, I feel that this, too, leads to confusion, because there is a need to distinguish between the awareness which all sentient creatures seem to have and the self-

awareness that only humans seem to possess. For example, a cat stalking a bird is very aware of the bird and the desired outcome, but does not possess the self-awareness of its owner, who remembers how she felt when she found the last dead and mangled bird lying at the back door. Psychology in particular has a difficulty in recognising that awareness is not the same as self-awareness, and that consciousness includes, but is not reducible to, experience. Thus, awareness is junior to consciousness, whereas self-awareness seems (at least in part) to characterise consciousness.

3. Consciousness and cognition: The term *cognition* is used by most psychology texts in a narrower sense than that of consciousness, and I regard it as hierarchically junior to consciousness. The Latin origin (*cognitio*) shows that its root meaning is *to become acquainted* with something, or to *learn* something, or to *acquire knowledge*. That it is used as a lesser term than consciousness is evident in the way that many (if not most) cognitive scientists use it (Gardner, 1985), where they exclude altered states and the affects, leaving cognition as embracing all those processes that are commonly regarded as employing the mind (and the mind's representations) to do the work (eg, memory and thinking). Thus, in the sense in which cognitive science employs it, cognition appears to be only **one** of the operations of consciousness.

4. Consciousness and the brain: It is difficult to consider the topic of consciousness without also considering its physical organ of expression -- the brain. It is the difficulty of considering consciousness apart from a physical organ which makes it so easy for many consciousness theorists to lean toward a physicalist-reductionist stance. Even where the view is not one of extreme materialism, consciousness is still regarded as dependent in some way or other upon the brain. The theoretical views on this issue range from seeing consciousness simply as a byproduct of the neural activity in the brain, to seeing consciousness as having an existence quite independent of the brain even though it must express itself through a brain. During the course of my thesis, I shall show that consciousness is as much an explicate of the implicate order as is the brain.

At this point a brief digression is necessary in regard to dualism in general and Descartes' version in particular. The model I shall develop in the course of this thesis depends on two quite distinct entities -- the implicate order and the explicate order. The fact that these two orders have very different property lists and are not reducible one to another shows that my model is essentially dualistic. Descartes developed a dualistic system in making such a clear and sharp distinction between mind and body (see the earlier discussion). His distinction was so sharp and rigid that there arose the problem as to how two such totally different categories of thing could ever interact. From this Cartesian problem or trap, there arose the notion that all dualistic schemes have an interaction problem. I do not believe this to be the case, and I shall show (in Chapter 3) that my dualism does not suffer from the Cartesian trap, and that not only is interaction between the implicate and explicate orders possible, but is necessary to the process of evolution.

Also, in the model I develop in Chapter 3, I shall show that consciousness is the explication of a very high region within the implicate order (that I have called Mind), where this explication is the result of a complex interaction between evolving neural structures and the implicate order. In that consciousness is as much an explicate as is the brain, there is no dualism and hence no issue in regard to the Cartesian trap. Interaction between consciousness and brain can and does occur as between so many other aspects of the explicate order (eg, between the software and hardware in a computer).

I will not be able to define consciousness with any precision. This definition will come out of the expositions in Chapters 3 and 4. However, it is worth setting out now some of the key aspects of what we call consciousness, so as to place these on some conceptual map for future use. To ease this task, I will deal at present only with human consciousness, wherein certain key issues arise.

Firstly, there is implied an organisation which is interposed between the vegetative life of an organism and the world with which it is in relation.

Secondly, consciousness appears to objectify itself and reflects itself in a model of its world. That is, consciousness is in the world because the world (especially that of others) enters into its constitution.

Thirdly, the construction of a model of the objective world appears to be the task of consciousness and its being.

Fourthly, consciousness organises itself either simultaneously (synchronistically) or historically (diachronistically) with reference to sensory data.

Finally, consciousness is organised so as to have an experience at each moment of its history, and to manifest as the person which emerges through this history.

Long ago, (in 1690) John Locke (Locke, 1961) listed many constituents of consciousness, such as perception, thinking, doubting, believing, reasoning, knowing and willing. Any comprehensive definition must include criteria for the structure and function of each constituting agency. Borrowing somewhat from Locke's original and thoughtful analysis, human consciousness appears to be strongly related to: **affective life; experiencing reality; attention-reflection; personality** and to **volition** (will). Therefore, any attempt at defining human consciousness entails considering each of these factors in turn, and then considering the meta-nature of that which organises and relates them.

Consciousness and the affective life: Far from defining consciousness, the affective state presupposes consciousness to be the very condition of experience. That is, to be conscious is to have sensations, all of which affect the body or set off reactions within it. This is another

way of saying that to be conscious is to experience sensations (internal and external) and, as a result, to feel.

Consciousness and the experience of reality: Conscious beings perceive an apparent reality, and in adapting themselves to this reality, unfold a complex operational capacity. To be conscious is to know one's experience. That is, to be conscious is to be capable of grasping one's own knowledge in the categories of verbal communication. This must be true even for those sense modes (eg, olfactory) for which we lack a sophisticated vocabulary, by comparison with, say, that of vision.

The major implication here is that the human mode of consciousness is, to a large extent, defined in terms of the speech (internal or external) capacity. However, this ignores that which we call unconscious, in that what is said (internally or externally) seems to sit on some deep structure of language that is never present to consciousness. We are aware only of segments of a given train of thought, as occur in a conversation, or as I type these words, and are never aware of the entire train. It is only when we have got that train out (verbally or otherwise) can we see it as a complete entity, and then realise that prior to this, the greater bulk of it was unconscious.

Attention-Reflection: Attention expresses the notion of a tension toward some desired goal, where the degrees of attention are hierarchical, ranging from involuntary functions to free-creative acts. Alertness, attention and wakefulness are synonyms, and are the result of a dynamic function and structure, which arrives at its optimal power of differentiation only by acquainting itself with the infrastructures from which it emerges. This is a key notion, and lies at the root of the evolution of consciousness, in that consciousness evolves to the extent that the enfolded infrastructure becomes a part of the awareness of a conscious being.

Reflection is the process by which thought returns in on itself and duplicates the acts in the external world which it directs. It is carried to its furthest power by attention. This might be

understood to mean that reflection is some higher mode of consciousness and not synonymous with ordinary waking consciousness. Perhaps this is where consciousness gets closest to pure ideation, which makes the assumption that there is a realm of pure ideas to operate in. What I am trying to convey is that consciousness is not a given state or particular experience, but is that meta-structure within which there are a variety of modes, which seem to be arranged hierarchically, with reflection being a higher state than, say, a cognitive process such as numerical reasoning.

Personality: The self is complex, and not simple, where personality is a history, linking the self's modes into a series of events, and the self is the author of its own person. If self-consciousness involves founding one's own person, and if conscious being is the very nature of the person, then neither the totality nor the basic structure of the person can be reduced to this manner of being conscious or to this idea of consciousness. Personality and self represent the transcendental aspect of *being someone* with respect to that person *having consciousness of some thing*.

Volition (Will): Philosophically, the issue of volition or will has been linked with that of the moral or ethical sense. This leads into the view that consciousness becomes moral consciousness when it evaluates-reflects upon its values, where moral consciousness cannot be radically separated from psychological consciousness. Morality is far from being some absolute entity or state, because it is caught up in that sea of reflections called consciousness.

In considering affects, experiencing, attention-reflection, personality and volition as the major factors of consciousness, it would be easy to fall into the trap of assuming that these factors in combination comprise consciousness. That is, to regard consciousness as simply the sum of a collection of parts. Conversely, it would be just as easy to assume that consciousness is some diffuse thing that permeates its various psychical structures, thus refusing to consciousness its own structural integrity. In the first view there is an indifference

to the interconnections of the parts and to any meta-aspect. The second view seems to oppose the rooting of consciousness in the body, such that consciousness is not regarded as a natural phenomenon because it is utterly transcendent. It is my view that neither of these approaches is correct, because I regard consciousness as existent in its own right, as having a hierarchical order and as being the meta structure that integrates the parts without being the sum of them, and as being expressible through a physical form.

However, note from my earlier argument that, in positing consciousness as an explicate in its own right, I am avoiding the kind of dualism that Descartes created, with its attendant difficulties. Rather, consciousness is a very high order explicate of the implicate order. Thus, in my thesis, consciousness (mind in common parlance) and brains (bodies) are both explicates, hence there is no dualism, and no problems regarding their interaction. I have already conceded that my scheme is dualistic in that there is an implicate order and an explicate order. However, as will be shown in Chapter 2 (and reinforced in Chapter 3), the explicate order derives from the implicate order, where interaction is not only feasible but essential to the evolutionary process.

THE IMPLICATE-EXPLICATE SET OF CATEGORIES

As mentioned earlier, the terms "implicate" and "explicate" arise out of the work done by Bohm (1980), who sought to explain the fundamental paradoxes which emerge out of experiments carried out at the quantum level of matter.

The problems encountered by researchers at the quantum mechanical level will be dealt with more fully in Chapter 2. However, in simplified form, the most profound of the quantum paradoxes, and the root mystery, is the wave-particle duality of light or other subatomic phenomena such as electrons. That is, under certain conditions light will behave as a wave and thus be capable of producing interference patterns. Under other conditions light will behave in a corpuscular fashion, as individual photons, and so will not produce interference

patterns. To explain this paradox and yet remain within the agreed quantum mechanical theory, David Bohm and co-workers (eg, Bohm & Bub, 1966) postulated the existence of hidden variables. The subsequent theorising which arose out of this proposal led Bohm to postulate an implicate order, which he more fully developed several years later (Bohm; 1973; 1980).

Bohm's notions will be dealt with more fully in Chapter 2. However, because Bohm's notion of an implicate order is crucial to my thesis, I need to dwell briefly on it here. In essence, the implicate order is intangible and so non-physical, and yet contains enfolded within itself all that we regard as physical matter, which is an explication of what is enfolded. In this view, matter is an explicate of an implicate order, and as such is not the solid substantial tangible stuff beloved of physicalists. All that we regard as matter, which includes brains, has its laws and being in an implicate order. I shall argue that consciousness is an explication of something that is intangible and is, in my view, the realm of Mind (a very high region within the implicate order).

Where the physicalist insists on the identity *mind = matter*, and says that what the equals sign means is that mind (consciousness in my terminology) and matter are properties of the same thing, I wholly agree. That is, mind or consciousness is an explicate, just as is matter (eg, brains). Thus they are of the same order of things, in that, as explicates, they are both derivatives of the implicate order. However, I shall show that consciousness is not reducible to matter (brains), because it is an explicate of a very high region within the implicate order (Mind), where the brain (as an explicate) derives from a lower region within the implicate order. Mind (with a capital M) is wholly implicate (is never explicated) and consciousness (as a very high order of explication of Mind) acts somewhat as the agent of the implicate realm. In this context, brains are purely explicate although of a very complex nature.

I will demonstrate that, although consciousness is an explicate entity, it is a very different an autonomous entity from that of the brain, and is not reducible to matter per se. But more than

this, I will attempt to show that there is no dualism at the level of consciousness-brains. In this view, brains (as much as stones and stars) are an explication of an intangible implicate order, which contains within itself all that we perceive as law-like and displaying order. Beyond this, I will attempt to show that evolution per se, and in particular the evolution of consciousness, is the result of the interaction that takes places between the implicate and explicate realms. Note that in arguing thus, I am not saying that the explicate realm is unreal or ghostly. It is real enough to our sensory apparatus. How could it be otherwise, when that very sensory apparatus is itself of the explicate order.

THE EVOLUTION OF CONSCIOUSNESS

We are now in a better position to understand what is meant by the combination of the two key terms of this thesis: *evolution* and *consciousness*. In a highly simplified form, the evolution of consciousness is that process wherein the implicate order unfolds its potential (as Mind) and in so doing produces at the explicate level organisms having a self-reflecting and self-willing awareness.

If we take a strictly Darwinian view, the term **evolution of consciousness** has little meaning because consciousness per se has little place in a scheme that is so blatantly physicalist-reductionist. As seen earlier, Darwin's world view is based on a chance-driven mechanism, and evolution relates to the differentiation of genetic material in populations. The Lamarkian notion of the individual influencing its own genetic code is rejected by the Darwinian view, hence individual consciousness plays no role in the evolutionary process. In fact, the Darwinian view seems to counter the original meaning of the word evolution. There is no unfolding of what was originally enfolded. There is simply a biological progression from the very simple to the very complex, wherein chance and adaptation work together, and the unit of evolution is a species.

For these reasons, while I recognise the existence of Darwin's views, I shall not be particularly constrained by them. Although I accept the usage of the term evolution as applied to biological forms, I widen its usage to embrace the topic of the evolution of consciousness. This stance is based on the declared belief that Mind (hence its derivative consciousness) is not reducible to matter or even energy, but is (in the sense that Bohm uses the term implicate), the basis or substratum of that which is perceived as matter-energy. In this view, consciousness is the condition that arises in the interaction between the implicate order and its explication in matter-energy. Consciousness varies in degree, according to its stage of evolution. In the implicate order, consciousness is a potential only. It is in the explicate order that consciousness becomes discrete and varies in the degree of its manifestation.

-- oOo --